Technical Report No.144

23

SACLANT ASW

RESEARCH CENTRE

A PROBABILISTIC THEORY OF ANTI-SUBMARINE WARFARE MODELS DEVELOPED IN TERMS OF CONGESTION THEORY

by

BRIAN W. CONOLLY

15 APRIL 1969

NATO
VIALE SAN BARTOLOMEO, 400
I-19026-LA SPEZIA, ITALY

This document is released to a NATO Government at the direction of the SACLANTCEN subject to the following conditions:

1. The recipient NATO Government agrees to use its best endeavours to ensure that the information herein disclosed, whether or not it bears a security classification, is not dealt with in any manner (a) contrary to the intent of the provisions of the Charter of the Centre, or (b) prejudicial to the rights of the owner thereof to obtain patent, copyright, or other like statutory protection therefor.

2. If the technical information was originally released to the Centre by a NATO Government subject to restrictions clearly marked on this document the recipient NATO Government agrees to use its best endeavours to abide by the terms of the restrictions so imposed by the releasing Government.

TECHNICAL REPORT NO. 144

SACLANT ASW RESEARCH CENTRE
Viale San Bartolomeo 400
I 19026 - La Spezia, Italy

A PROBABILISTIC THEORY OF ANTI-SUBMARINE WARFARE MODELS DEVELOPED IN TERMS OF CONGESTION THEORY.

By

Brian W. Conolly

15 April 1969

APPROVED FOR DISTRIBUTION

Ir M.W. van Batenburg
Director

Manuscript Completed:
24 February 1969

TABLE OF CONTENTS

A PROBABILISTIC THEORY OF ANTI-SUBMARINE WARFARE MODELS DEVELOPED IN TERMS OF CONGESTION THEORY

By

Brian W. Conolly

ABSTRACT

This report, which is methodological, develops a probabilistic theory that has direct application to both antisubmarine warfare and congestion models. The theory is expressed in congestion terminology because of the presumed wider knowledge and appeal of this field. This results in a simplified presentation of the general theory of infinite service facility systems with specific application to $M/Y/\infty$ and $X/M/\infty$, some of which have already been studied by Takács and Khintchine. A new result is given for the output of the latter process. The analogy between certain infinite service facility systems and a single-server system with queue length dependent service is exploited to provide results for the latter process. A further new result for the busy period of such a process is quoted. The antisubmarine applications are to the formally similar models of the number of units present in a geographical area, and to the attrition of an enemy submarine force subjected to a steady threat from an antisubmarine barrier that geographical or other constraints compel it to transit.

INTRODUCTION

This report is concerned with the theory of a probabilistic model that may be used to describe generalized situations that arise in a military context, in particular in anti-submarine warfare. One situation is as follows: Units arrive at the boundary of a geographical area in which they propose to spend some time. They may be ships about to cross an ocean or strait, singly or in convoy; or they may be submarines going on patrol. In a non-military situation they may be cars entering a car park, or prospective clients arriving at a supermarket. As long as the units are not prevented from achieving this objective they may be described in the language of Congestion or Queueing Theory as "customers" entering a service facility with infinite capacity, that is to say, one that can provide service to any unit demanding it.

In these situations one needs to make statements about the total number of units in the area at any given time, and about the number of units that leave the area in a given time interval. These are respectively the content of the area, and the output from the area. In a defensive situation one might be interested in the number of merchant ships in an area subjected to a submarine, or other, threat simply because one wants to know the scale of protection it is required to provide. From the same point of view, one is likely to want to know the magnitude of an enemy threat in the area in order to be able to assess the level of forces required to subdue it. The output from an area is equally of interest, particularly if the "service time" within the area includes the possibility of destruction as well as safe transit.

Numerical statements about content and output are most easily made in deterministic terms. Thus, suppose the content of an area at time t is $\xi(t)$ units, while the rate of arrival at the boundary is λ units per day and the mean time required by each unit to transit the area is μ^{-1} days. Then the mean number of units leaving the area per day will be $\mu\xi$ and we shall have

$$\frac{d\xi}{dt} = \lambda - \mu\xi\,, \qquad \text{[Eq. 1]}$$

so that if λ and μ are time independent and $\xi(0) = \xi_0$,

$$\xi(t) = \xi_0\, e^{-\mu t} + \rho(1 - e^{-\mu t}), \qquad \text{[Eq. 2]}$$

where $\rho = \lambda/\mu$. The output $\eta(T)$ from the area in the interval $(0, T)$ is then

$$\eta(T) = \mu \int_0^T \xi(t)\, dt$$

$$= \xi_0(1 - e^{-\mu T}) + \rho\, f(T) \qquad \text{[Eq. 3]}$$

where

$$f(T) = \mu T - (1 - e^{-\mu T}).$$

Predictions made on the basis of such a theory can be useful as indicators, but what is really needed in a more operationally

realistic treatment is a description of $\xi(t)$ and $\eta(t)$ in probabilistic terms. This can be achieved by studying the basic quantities

$$p_n(t) = \Pr\left[\xi(t) = n\right]$$

and

$$v_n(t) = \Pr\left[\eta(t) = n\right]$$

with suitable initial conditions. These are not, however, the only probabilities of interest in an operational study. The duration of time, for example, when $\xi(t) > N$, where N is some prescribed number, is also of interest, both in relation to targets and to the threat.

The same model can be used to describe the losses to an enemy force of submarines that, in order to operate, has to cross an antisubmarine barrier where a constant kill rate per transit can be applied. Let $\xi(t)$ be the number of enemy submarines on operational strength at time t, and λ the building rate. If all submarines are used, then the rate of arrival at the barrier is proportional to $\xi(t)$, and the number destroyed per unit time can be written $\mu\xi(t)$. Thus, deterministically,

$$\frac{d\xi}{dt} = \lambda - \mu\xi ,$$

which is the same formal model as that given in Eq. 1 and the same probabilistic theory will apply.

It will be noticed that the deterministic differential equation given above is suggestive of radioactive decay. Familiarity with warfare models of the Lanchester type and with epidemic and struggle-for-survival models in ecology will lead to the recognition that the model belongs also to this family, though with less complexity in that we are not dealing directly with the interaction of two species. More generally the model is a particular kind of birth and death process.

In the language of Queueing Theory the problem of making a probabilistic description of the content $\xi(t)$ of a geographical area is known as Erlang's Problem for an Infinite Collection. Erlang was specifically concerned with the application to telephone exchanges and in this context $\xi(t)$ is the number of calls in progress at time t. Since Queueing Theory and its terminology are widely known it seems appropriate to proceed with the development of the theory in the queueing theoretic context, not forgetting the wide range of applications in other spheres of activity. In the terminology of queueing theory, "units" are referred to as "clients" or "customers". Their "service time" is the time spent in the "service facility" or geographic area. The "input" to the "system" consists of a description of the intervals of time separating the arrivals of successive customers/units at the service facility. A principal objective of this report is a simplified exposition of the probabilistic theory of infinite capacity service systems that are formally identical to the military model discussed in the previous paragraphs.

A mathematical description of service systems requires a specification in probabilistic terms, where appropriate, of the input, the service time, in some cases the "queueing discipline" (the order in which customers are served, for example), as well as of the number of service points. On the basis of these one makes probabilistic statements about how long a typical customer has to wait, the length of the uninterrupted periods during which the facility is kept busy, the interval of time between successive departures from the

system, and so on. To these may be added Erlang's Problem, which is also of particular interest to the designers of supermarkets and doctors' waiting rooms, as well as to those whose task it may be to protect shipping against enemy submarines about which the only thing known is a description of the input and cycle time in probabilistic terms.

Erlang's Problem for a finite collection is rather intractable, particularly where so-called exact finite time results are concerned. These are required when one wants to make statements about $\xi(t)$, which it will be recalled is the total number of units present at time t after the system was initiated, rather than about how $\xi(t)$ behaves in the long run. It turns out, on the other hand, perhaps at first sight surprisingly, that finite time results are available for certain systems that offer infinite service potential and that are, at the same time, reasonable models for practical information about the upper limit that the system can achieve. It is therefore of obvious practical value.

Before proceeding further it is advantageous to introduce a fundamental notation of Queueing Theory. The terminology $X/Y/N$ is used to denote a service system with N servers. The units requiring service are supposed to arrive at time intervals whose distribution is denoted by X. If the service facility is full, they wait. When they reach a service point the time required to complete service has a distribution specified by Y. The most commonly encountered type of X and Y is the negative exponential distribution. This is always denoted by the letter M. Another type of frequent occurrence relates to the deterministic case, and is always denoted by D. Thus $M/D/1$ denotes a single server system in which arrivals are separated by time intervals having a negative exponential distribution, and each customer receives a fixed service time. In this report we are dealing with systems that may be denoted by $X/Y/\infty$.

Explicit time-dependent solutions of Erlang's Problem for $M/Y/\infty$ and for a special $X/Y/\infty$ in which X is governed by a time dependent parameter, were given by Khintchine [Ref. 2]. Subsequently Takács [Ref. 3] gave a solution for $X/M/\infty$ where arrivals occur in a renewal process with arbitrary inter-arrival distribution. The Takács solution is expressed as a Laplace transform. Subsequently it was discovered [Mirasol (Ref.4)] that the output of the $M/Y/\infty$ system is Poisson. This aroused considerable interest since an infinite service system can be used to model delays in arrival at a service point [see Kendall (Ref. 5), for example]. That is to say, one may imagine that when a customer arrives he is put through an imaginary infinite server system before going to the serving point, and this has the effect of delaying him. The output of the infinite server system then becomes the input of the real service facility, and since queues with Poisson input (essentially of M/Y type) are well understood, Mirasol's discovery held out some promise of throwing light on the notoriously difficult delay problem. More recently Vere Jones [Ref. 6] has shown that the output of an infinite sequence of $X/Y/\infty$ systems, each feeding the other, is Poisson.

With the exception of a hint in Kendall [Ref. 5], no investigator seems to have exploited the fact that under suitable conditions the customers entering an infinite service facility may be thought of as flowing through it independently. This fact can be used to simplify considerably the development of the theory. It is the purpose of this report to demonstrate how this can be done, and to make an application to the waiting time problem for an interesting, incompletely investigated queueing system in which service is geared to demand.

1. THE METHOD

We suppose that at time $t = 0$ the infinite capacity service system opens its doors, permitting service to begin on the n customers who are waiting. First we give formulae for:

(a) The probability $a_{nm}(t)$ that m of the original n customers have been served by time t. This is clearly independent of subsequent arrivals who, because the service facility is infinite, will flow independently through the system.

(b) The probability density function $b_{nm}(t)$ that the m^{th} customer of the original n has been served in the small time interval $(t, t + dt)$.

Let the probability density function (p.d.f.) of the service time s of a customer be denoted by $b(t)$. We make no special demands on $b(t)$ except that

$$\int_0^\infty b(t)\, dt = 1$$

and that its moments exist. We denote the distribution function (d.f.) by $B(t)$, i.e.

$$B(t) = \int_0^t b(s)\, ds,$$

and the tail of the distribution by $B_c(t)$, so that

$$B_c(t) = 1 - B(t).$$

The m customers to have been served by time t can be selected in $\binom{n}{m}$ ways and since the probability that any of the original n has been served by time t is $B(t)$, we have immediately

$$a_{nm}(t) = \binom{n}{m} \left[B(t)\right]^m \left[B_c(t)\right]^{n-m}. \qquad \text{[Eq. 4]}$$

To deal with $b_{nm}(t)$ we note that the m^{th} customer who has been served in $(t, t+dt)$ can be chosen in n ways. The p.d.f. of his service time is $b(t)$. In addition we require that $m-1$ out of the remaining $n-1$ are served before t. Thus

$$b_{nm}(t) = n \binom{n-1}{m-1} \left[B(t)\right]^{m-1} \left[B_c(t)\right]^{n-m} b(t). \qquad \text{[Eq. 5]}$$

Writing D for $\frac{d}{dt}$, it is then clear that

$$b_{nm}(t) = \binom{n}{m} \left[B_c(t)\right]^{n-m} D\left[B(t)\right]^m. \qquad \text{[Eq. 6]}$$

Now we consider arrivals that occur subsequently to $t=0$. Again, because the service has infinite capacity, these can be dealt with independently of the fate of the original n customers in the system. We concern ourselves with

(a) <u>the probability</u> $r_{nm}(t)$ that during $(0, t)$

n arrivals take place, of which m have been served <u>before</u> time t, (not necessarily in order of arrival);

and

(b) <u>the associated probability density function</u> $q_{nm}(t)$ that in $(0,t)$ n arrivals take place, of which the m^{th} has been served in the small interval $(t, t+dt)$ (not necessarily in the order of arrival).

We denote the p.d.f. of the intervals between arrivals by $a(t)$, which, as for $b(t)$, we suppose to be a properly behaved function with moments. We also write

$$A(t) = \int_0^t a(s)\, ds .$$

and

$$A_c(t) = 1 - A(t).$$

Given that $t=0$ is an arrival instant (one might suppose that the latest of the customers waiting at $t=0$ arrived at that moment) we clearly have

$$r_{10}(t) = \int_0^t a(s)\, A_c(t-s)\, B_c(t-s)\, ds, \qquad \text{[Eq. 7]}$$

and

$$r_{11}(t) = \int_0^t a(s)\, A_c(t-s)\, B(t-s)\, ds. \qquad \text{[Eq. 8]}$$

Adding these we obtain, as might be expected, the probability of exactly one arrival in the interval $(0, t)$. The explanation of Eqs. 7 and 8 is as follows. We suppose that the one and only arrival occurs in the small time interval $(s, s+ds)$. The probability of this event is $a(s)\ A_c(t-s)\ ds$. In the remaining part of the interval, $t-s$, Eq. 7 requires that the arrival is not served, [probability $B_c(t-s)$], or in Eq. 8, is served, before epoch t [probability $B(t-s)$]. Adding over all possible s completes the explanation. Generalizing to $r_{nm}(1 \le m \le n-1)$ we have

$$r_{nm}(t) = \int_0^t a(s)\ B(t-s)\ r_{n-1,\, m-1}\ (t-s)\ ds +$$

$$+ \int_0^t a(s)\ B_c(t-s)\ r_{n-1,\, m}\ (t-s)\ ds. \qquad \text{[Eq. 9]}$$

The explanation is again based on a first arrival in the interval $(s, s+ds)$. Either this arrival is served before t or he is not. Invoking the independence of flow concept we see in the first case that during (s, t), $n-1$ arrivals must occur with $m-1$ services, and in the second, $n-1$ arrivals and m services. For $m=0$ and n we obtain single integrals:

$$r_{n0}(t) = \int_0^t a(s)\ B_c(t-s)\ r_{n-1,\, 0}(t-s)\ ds; \qquad \text{[Eq. 10]}$$

$$r_{nn}(t) = \int_0^t a(s)\ B(t-s)\ r_{n-1,\, n-1}(t-s)\ ds. \qquad \text{[Eq. 11]}$$

With the definitions

$$r_{00}(t) = A_c(t)$$

and

$$r_{nm}(t) \equiv 0$$

for $m > n$, and either m or n negative, the validity of Eq. 9 can be extended to all m and n, thus including Eqs. 7, 8, 10 and 11.

Turning to the corresponding densities $q_{nm}(t)$ we use a similar argument. This function has no meaning for $m = 0$. For $m = 1$ we obtain:

$$q_{n1}(t) = \int_0^t a(s) \left[b(t-s)\, r_{n-1,\,0}(t-s) + B_c(t-s) q_{n-1,\,1}(t-s) \right] ds; \qquad \text{[Eq. 12]}$$

$$q_{nm}(t) = \int_0^t a(s) \left[b(t-s)\, r_{n-1,\,m-1}(t-s) + B_c(t-s)\, q_{n-1,\,m}(t-s) + \right.$$

$$\left. + B(t-s) q_{n-1,\,m-1}(t-s) \right] ds; \qquad \text{[Eq. 13]}$$

for $2 \leq m \leq n-1$;

$$q_{nn}(t) = \int_0^t a(s) \left[b(t-s)\, r_{n-1,\,n-1}(t-s) + B(t-s) q_{n-1,n-1}(t-s) \right] ds. \qquad \text{[Eq. 14]}$$

The explanation of Eqs. 12, 13 and 14 is similar to that for Eqs. 9, 10 and 11. The first arrival occurs in $(s, s+ds)$ and we then have to consider whether or not he is the customer to be served in $(t, t+dt)$. In the case $m=1$ he either completes in $(t, t+dt)$ or after t; in the case $m=n$ he either completes in $(t, t+dt)$ or before t. For $2 \leq m \leq n-1$ he may complete in $(t, t+dt)$, before t or after t. There are no other possibilities.

Again a suitable definition permits the validity of Eq. 13 to be extended to all m, n. This is

$$q_{n,m}(t) \equiv 0$$

for $m=0$ and $m>n$.

It will be noted that subject to suitable conditions these formulae for the probabilities r_{nm} and densities q_{nm} are valid for arbitrary arrival and service intervals of recurrent type.

One can now turn to Erlang's Problem: that of writing down the probability

$$p_k^{(N)}(t) = \Pr\left[\xi(t) = k \mid \xi(0) = N\right],$$

where it will be recalled that $\xi(t)$ $(t \geq 0)$ is the non-negative, integer-valued stochastic process that describes the number of units in the system at time t after the process was initiated, and N is the number of units present at $t=0$. There are two ways of doing this and it will always be supposed that at $t=0$ all services begin simultaneously, and an arrival can be regarded as just having occurred.

First consider the case $0 \le k \le N$. It is then possible to achieve the condition $\xi(t) = k$ without any arrivals during $(0, t)$. The contribution to $p_k^{(N)}(t)$ is thus $A_c(t) a_{N,N-k}(t)$. When arrivals do occur it is still possible that the k remaining belong to the original N. The condition is that all n arrivals $(n \ge 1)$ are served before t. The contribution to $p_k^{(N)}(t)$ is thus $a_{N,\,N-k}(t) \sum_{n\ge1} r_{nn}(t)$.

Now consider that all but one of the arrivals in $(0, t)$ have been served before t. Then $k-1$ of the original customers must remain, and the contribution to $p_k^{(N)}(t)$ becomes $a_{N,\,N-k+1} \sum_{n\ge1} r_{n,\,n-1}(t)$. Generalizing, if s customers out of the new arrivals remain at t, then $k-s$ originals must remain too. In this case we have a contribution of $a_{N,\,N-k+s}(t) \sum_{n=s} r_{n,\,n-s}(t)$ while $0 \le s \le k$. This argument also depends on the independent flow concept. We have finally

$$p_k^{(N)}(t) = A_c(t)\, a_{N,\,N-k}(t) + \sum_{s=0}^{k} a_{N,\,N-k+s}(t) \sum_{n\ge\max(1,s)} r_{n,\,n-s}(t)$$

[Eq. 15]

for $0 \le k \le N$. With the definition $r_{00}(t) = A_c(t)$ note that the first term in Eq. 15 can be included in the sum. Thus Eq. 15 becomes

$$p_k^{(N)}(t) = \sum_{s=0}^{k} a_{N,\,N-k+s}(t) \sum_{n\ge s} r_{n,\,n-s}(t)\,. \qquad \text{[Eq. 15a]}$$

When $k \ge N+1$, arrivals must take place. For simplicity consider $k = N+1$. If none of the original N has been served before t then exactly one subsequent arrival must be

present at time t. The contribution to $p_{N+1}^{(N)}(t)$ is then $a_{N0}(t) \sum_{n \geq s+1} r_{n,\, n-s-1}(t)$. Thus

$$p_{N+1}^{(N)}(t) = \sum_{s=0}^{N} a_{Ns}(t) \sum_{n \geq s+1} r_{n,\, n-s-1}(t) . \qquad \text{[Eq. 16]}$$

It is not difficult to extend the argument to obtain the general formula

$$p_{N+k}^{(N)}(t) = \sum_{s=0}^{N} a_{Ns}(t) \sum_{n \geq s+k} r_{n,\, n-s-k}(t), \qquad \text{[Eq. 17]}$$

for $k \geq 1$.

Noting from the accepted definition of binomial coefficients that

$$\binom{n}{m} \equiv 0$$

for $m > n$ and negative n or m, we then have automatically

$$a_{nm}(t) \equiv 0$$

for such values of m and n, and thus the validity of Eq. 15a can be extended to all $k (\geq 0)$. This provides the solution to Erlang's Problem in some generality. The argument, though correct, is rather dull and the formulae on their own are not very exciting.

A more aesthetically attractive formulation, which emphasizes the independence of flow concept, leads to the following set of equations for $p_r^{(N)}(t)$:

$$p_{N-k}^{(N)}(t) = a_{Nk}(t)\ A_c(t) + \sum_{n=0}^{N-k} a_{N,\,k+n}(t) \int_0^t a(s)\ p_n^{(1)}(t-s)\ ds, \qquad \text{[Eq. 18]}$$

for $0 \le k \le N$;

$$p_{N+k}^{(N)}(t) = \sum_{n=0}^{N} a_{N,\,n}(t) \int_0^t a(s)\ p_{k+n}^{(1)}(t-s)\ ds, \qquad \text{[Eq. 19]}$$

for $0 < k < \infty$.

We consider first Eq. 18. Here the number of units present at time t is at most equal to N, the number originally present. There may have been no arrivals, in which case provision has to be made for the right number of departures from among the initial N. This accounts for the first term on the right hand side. The second term deals with the case where arrivals have taken place. We now argue that because the service facility has infinite capacity, the first arrival creates a situation in which a second infinite service facility is initiated with a single unit present. We then have to provide that the sum of the number remaining at time t out of the initial batch of N, plus the number remaining in the second system initiated with a single unit at time s, has the correct value.

When the number present at time t exceeds the initial number N, arrivals have to occur. The explanation of the single term on the right hand side of Eq. 19 is then identical with that of the second term on the right hand side of Eq. 18.

A number of steps can be taken to verify the correctness of these general formulations of the solution of Erlang's Problem for an infinite service facility. One can, for instance, show that $\sum_{k\geq 0} p_k^{(N)} = 1$, as it should, for all $N \geq 1$. In addition one can demonstrate the equivalence of Eq. 15 with Eqs. 17, 18 and 19 by direct substitution. It would be tedious to continue in full generality and we shall therefore proceed in the framework of particular systems.

It seems worth pointing out that the theory permits one to handle units that require different grades of service. Suppose for example that the input consists of a variety of units, some having inter-arrival interval density function $a_1(t)$, others $a_2(t)$, and so on, and that service on the $a_1(t)$ stream has density function $b_1(t)$, $b_2(t)$ on the $a_2(t)$ stream, etc. Provided that $t = 0$ marks an arrival and beginning of service epoch for each stream of units, the different types can be allowed to flow independently through the infinite system. The number $\xi(t)$ of units present at time t is then, with obvious notation, given by

$$\xi(t) = \xi_1(t) + \xi_2(t) + \ldots\ldots$$

and $\Pr\left[\xi(t) = n\right]$ is the convolution of the individual probabilities given by the theory. Thus one can with the same theory describe, for example, a mixture of slow and fast ships crossing an area.

A number of steps can be taken to verify the correctness of these general formulations of the solution of Erlang's Problem for an infinite service facility. One can, for instance, show that $\sum_{k\geq 0} p_k^{(N)} = 1$, as it should, for all $N \geq 1$. In addition one can demonstrate the equivalence of Eq. 15 with Eqs. 17, 18 and 19 by direct substitution. It would be tedious to continue in full generality and we shall therefore proceed in the framework of particular systems.

It seems worth pointing out that the theory permits one to handle units that require different grades of service. Suppose for example that the input consists of a variety of units, some having inter-arrival interval density function $a_1(t)$, others $a_2(t)$, and so on, and that service on the $a_1(t)$ stream has density function $b_1(t)$, $b_2(t)$ on the $a_2(t)$ stream, etc. Provided that $t = 0$ marks an arrival and beginning of service epoch for each stream of units, the different types can be allowed to flow independently through the infinite system. The number $\xi(t)$ of units present at time t is then, with obvious notation, given by

$$\xi(t) = \xi_1(t) + \xi_2(t) + \ldots\ldots$$

and $\Pr\left[\xi(t) = n\right]$ is the convolution of the individual probabilities given by the theory. Thus one can with the same theory describe, for example, a mixture of slow and fast ships crossing an area.

2. THE SYSTEM M/Y/∞

2.1 Erlang's Problem

This is the system for which the solution of Erlang's Problem was first furnished by Khintchine [Ref. 2], whose results are given here. The arrival pattern is governed by a negative exponential distribution with parameter λ. That is to say

$$a(t) = \lambda e^{-\lambda t}, \qquad \text{[Eq. 20]}$$

and

$$A(t) = 1 - e^{-\lambda t},$$

$$A_c(t) = e^{-\lambda t}. \qquad \text{[Eq. 21]}$$

Y denotes a general service time distribution whose p.d.f. and distribution function will continue to be denoted by $b(t)$, $B(t)$. The simplest approach to the probabilities $p_n^{(N)}(t)$ is furnished in this case by the functions $r_{nm}(t)$ [c.f. Eqs. 9, 10 and 11]. Clearly,

$$r_{10}(t) = \lambda \int_0^t e^{-\lambda s}\, e^{-\lambda(t-s)}\, B_c(t-s)\, ds$$

$$= \lambda\, e^{-\lambda t} \int_0^t B_c(t-s)\, ds$$

$$= \lambda\, e^{-\lambda t}\, g(t)\,, \qquad \text{[Eq. 22]}$$

where

$$g(t) = \int_0^t B_c(s)\, ds . \qquad \text{[Eq. 23]}$$

Similarly,

$$r_{11} = \lambda \int_0^t e^{-\lambda s}\, e^{-\lambda(t-s)}\, B(t-s)\, ds$$

$$= \lambda\, e^{-\lambda t} f(t) , \qquad \text{[Eq. 24]}$$

where

$$f(t) = \int_0^t B(s)\, ds. \qquad \text{[Eq. 25]}$$

It will be noticed that $f(t) + g(t) = t.$ [Eq. 25a]

Then it is easy to show inductively that

$$r_{nm}(t) = \lambda^n\, e^{-\lambda t}\, \frac{f^m(t)\ g^{n-m}(t)}{m!\ (n-m)!} \qquad \text{[Eq. 26]}$$

and that

$$q_{nm}(t) = \frac{\lambda^n\ e^{-\lambda t}\ g^{n-m}(t)}{m!\ (n-m)!}\ D\left[f^m(t)\right] \qquad \text{[Eq. 27]}$$

where $D \equiv \frac{d}{dt}$. Then from Eqs. 15 and 17 we obtain

$$p_k^{(N)}(t) = e^{-\lambda g(t)} \sum_{m=\max(0,k-N)}^{k} \binom{N}{N-k+m} B^{N-k+m}(t)\, B_c^{k-m}\, \frac{\left[\lambda g(t)\right]^m}{m!}$$

[Eq. 28]

for all $k \geq 0$, a weighted combination of compound Poisson terms with time-dependent parameter $\lambda g(t)$. If we put $N = 1$ we have, omitting arguments,

$$p_0^{(1)} = e^{-\lambda g} \; B \, , \qquad \text{[Eq. 29a]}$$

$$p_0^{(1)} = e^{-\lambda g} \left[B_c + B \, \frac{(\lambda g)}{1!} \right] , \qquad \text{[Eq. 29b]}$$

$$p_2^{(1)} = e^{-\lambda g} \left[B_c \, \frac{(\lambda g)}{1!} + B \, \frac{(\lambda g)^2}{2!} \right] , \qquad \text{[Eq. 29c]}$$

$$p_n^{(1)} = e^{-\lambda g} \left[B_c \, \frac{(\lambda g)^{n-1}}{(n-1)\,!} + B \, \frac{(\lambda g)^n}{n!} \right] . \qquad \text{[Eq. 29d]}$$

Clearly $\sum_{n \geq 0} p_n^{(1)} = 1$.

The special forms of Eqs. 18 and 19 for $N = 1$ are, since $a(t) = \lambda e^{-\lambda t}$:

$$p_0^{(1)} = e^{-\lambda t} \, B + \lambda B \, e^{-\lambda t} \int_0^t e^{\lambda s} \, p_0^{(1)}(s) \, ds; \qquad \text{[Eq. 30a]}$$

$$p_1^{(1)} = e^{-\lambda t} \, B_c + \lambda B_c \, e^{-\lambda t} \int_0^t e^{\lambda s} \, p_0^{(1)}(s) ds +$$

$$+ \lambda B \, e^{-\lambda t} \int_0^t e^{\lambda s} \, p_1^{(1)}(s) \, ds; \qquad \text{[Eq. 30b]}$$

$$p_{k+1}^{(1)} = \lambda B_c \, e^{-\lambda t} \int_0^t e^{\lambda s} \, p_k^{(1)}(s) ds + \lambda B \, e^{-\lambda t} \int_0^t e^{\lambda s} \, p_{k+1}^{(1)}(s) \, ds.$$

[Eq. 30c]

Substituting Eqs. 29a in the integral in Eq. 30a gives

$$\int_0^t e^{\lambda s - \lambda g(s)} B(s)\, ds = \int_0^t e^{\lambda f(s)}\, df(s) = \frac{1}{\lambda}\left[e^{\lambda f(t)} - 1\right]$$

whence it is clear that Eq. 30a is satisfied by Eq. 29a. Similarly, in the general case. The coefficient of $\lambda B\, e^{-\lambda t}$ in Eq. 30c is

$$\int_0^t e^{\lambda f(s)} \left[B_c(s) \frac{\left[\lambda g(s)\right]^k}{k!} + B(s) \frac{\left[\lambda g(s)\right]^{k+1}}{(k+1)!}\right] ds$$

$$= \frac{1}{\lambda} \int_0^t d\left[e^{\lambda f(s)} \frac{\left[\lambda g(s)\right]^{k+1}}{(k+1)!}\right]$$

$$= \frac{1}{\lambda} \left[\left(e^{\lambda f(s)} \frac{\left[\lambda g(s)\right]^{k+1}}{(k+1)!}\right)_0^t\right]$$

$$= \frac{1}{\lambda} e^{\lambda f(t)} \frac{\left[\lambda g(t)\right]^{k+1}}{(k+1)!} .$$

Thus the right hand side of Eq. 30c becomes

$$B_c\, e^{-\lambda g} \frac{(\lambda g)^k}{k!} + B\, e^{-\lambda g} \frac{(\lambda g)^{k+1}}{(k+1)!} ,$$

which is $p_{k+1}^{(1)}$.

2.2 The Output

Description of the output process from an infinite service facility can be quite general, as we shall now briefly indicate. Let $\eta(t)$ be the number output from an infinite service facility in the interval $(0, t)$ and let

$$v_k^{(n)}(t) = \Pr\left[\eta(t) = k \mid \xi(0) = n\right], \qquad \text{[Eq. 31]}$$

on the assumption that $t = 0$ corresponds both to the beginning of an arrival interval and to the beginning of service of all units. Then for $k = 0$ we have either

$$v_o^{(n)}(t) = A_c(t)\; a_{n0}(t) + a_{n0}(t) \int_o^t a(s)\; v_o^{(1)}(t-s)\; ds\,, \qquad \text{[Eq. 32]}$$

or

$$v_o^{(n)}(t) = A_c(t)\; a_{n0}(t) + a_{n0}(t) \sum_{m \geq 1} r_{m0}(t). \qquad \text{[Eq. 33]}$$

For the $M/Y/\infty$ system with $a(t) = \lambda e^{-\lambda t}$, Eq. 33 gives

$$v_o^{(n)}(t) = B_c^n(t)\; e^{-\lambda f(t)}\,. \qquad \text{[Eq. 34]}$$

Then

$$\int_0^t a(t-s)\, v_0^{(1)}(s)\, ds = \lambda e^{-\lambda t} \int_0^t e^{\lambda s} \quad B_c(s)\, e^{-\lambda f(s)}\, ds$$

$$= \lambda e^{-\lambda t} \int_0^t e^{\lambda g(s)}\, dg(s)$$

$$= e^{-\lambda t}(e^{\lambda g(t)} - 1)\,,$$

and it is clear that Eq. 32 is satisfied. The argument leading to Eqs. 32 and 33 is that none of the original n must have been served in $(0, t)$, nor must any of the subsequent arrivals. In the case of formulation by Eq. 32 it is argued that the first arrival initiates the activity of a second infinite service facility. The reader will readily verify that a general formulation on the lines of Eq. 32 gives:

$$v_0^{(n)}(t) = A_c(t)\, a_{nk}(t) + \sum_{i=0}^{k} a_{ni}(t) \int_0^t a(t-s)\, v_{k-i}^{(1)}\,(s)\, ds,$$

[Eq. 35a]

for $k \leq n$; and

$$v_k^{(n)}(t) = \sum_{i=0}^{n} a_{ni}(t) \int_0^t a(t-s)\, v_{k-i}^{(1)}\,(s)\, ds \qquad \text{[Eq. 35b]}$$

for $k > n$.

In the spirit of Eq. 33 we have

$$v_k^{(n)}(t) = A_c(t)\, a_{nk}(t) + \sum_{i=0}^{k} a_{ni}(t) \sum_{m \geq \max(1,k-i)} r_{m,k-i}(t),$$

[Eq. 36a]

for $k \leq n$; and

$$v_k^{(n)}(t) = \sum_{i=0}^{n} a_{ni}(t) \sum_{m \geq k-i} r_{m,\, k-i} \quad , \qquad \text{[Eq. 36b]}$$

for $k > n$.

For the system $M/Y/\infty$ these reduce to

$$v_k^{(n)}(t) = e^{-\lambda f(t)} \sum_{i=0}^{\min(k,n)} a_{ni}(t) \frac{\left[\lambda f(t)\right]^{k-i}}{(k-i)!} \qquad \text{[Eq. 37]}$$

for all $k \geq 0$. Thus, as for the state probabilities, the number output from the system in $(0, t)$ has for distribution a weighted sum of compound Poisson terms, but with time dependent parameter $\lambda f(t)$. This is essentially the finding of Mirasol [Ref. 4].

3. THE SYSTEM $X/M/\infty$

3.1 Erlang's Problem

Although the integral equation formulation for the output and state process seems more elegant it is not clear that it is as suggestive of the solutions as the alternative expressed in terms of the functions $r_{nm}(t)$, at least so far as the system $M/Y/\infty$ is concerned. The reverse is the case for the system $X/M/\infty$ and, by contrast with $M/Y/\infty$, it seems expedient to resort to the Laplace transformation. The density function $a(t)$ is now general, whereas $b(t)$ has the specific form $\mu e^{-\mu t}$.

The solution of Erlang's Problem has been published, with misprints, by Takács [Ref. 3]. It requires much algebra, which it would be tedious to reproduce. We content ourselves with a statement of the result for $\xi(0) = 1$ together with an indication of a method for its derivation.

We drop the superscript on $p_k^{(n)}(t)$, it now being understood that it is unity. We also write

$$\pi_k(z) = \int_0^\infty e^{-zt} p_k(t)\, dt$$

for the Laplace transform of $p_k(t)$, and further introduce the notation

$$\left.\begin{aligned} \alpha_k(z) &= \int_0^\infty e^{-(z+k\mu)t}\, a(t)\, dt, \\ \delta_k(z) &= 1 - \alpha_k(z), \\ f_k(z) &= \int_0^\infty e^{-(z+k\mu)t}\, A_c(t)\, dt; \end{aligned}\right\} \qquad \text{[Eq. 38]}$$

and henceforward drop explicit reference to the argument z. Then, for $k \geq 0$,

$$\pi_k = \sum_{i=0} \binom{k+i}{k} (-)^i \frac{\alpha_1 \alpha_2 \cdots \alpha_{k-1+i} \, f_{k+i}}{\delta_0 \delta_1 \cdots \delta_{k+i}} , \qquad \text{[Eq. 39]}$$

in which when $k=0$, the numerator is f_0 for $i=0$ and f_1 for $i=1$. The generating function

$$\Pi(x) = \sum_{k \geq 0} \pi_k x^k$$

can be written

$$\Pi(x) = \sum_{i \geq 0} (-)^i \frac{(1-x)^i \alpha_1 \alpha_2 \cdots \alpha_{k-1+i} \, f_{k+i}}{\delta_0 \delta_1 \cdots \delta_{k+i}} . \qquad \text{[Eq. 40]}$$

Putting $x=1$ gives

$$\Pi(1) = \frac{f_0}{\delta_0} = \frac{1}{z} ,$$

which implies that $\sum_{k \geq 0} p_k(t) = 1$, as it should.

The proof of Eq. 39 is carried out most expeditiously by introducing a shift operator E that is such that it increases the argument z of the various Laplace transforms by an amount μ. Thus

$$\alpha_k = E^k \alpha_0 , \qquad \text{[Eq. 41]}$$

for example. Later it will be convenient also to use a backward difference operator ∇ which is defined by

$$\nabla = 1 - E . \qquad \text{[Eq. 42]}$$

We notice also that the interpretation of expressions like $(1 - \alpha_0 E)^{-1}$ is

$$1 + \alpha_0 E + \alpha_0 E(\alpha_0 E) + \alpha_0 E(\alpha_0 E)(\alpha_0 E) + \ldots..$$

$$= 1 + \alpha_0 E + \alpha_0 \alpha_1 E^2 + \alpha_0 \alpha_1 \alpha_2 E^3 + \ldots.. \qquad \text{[Eq. 43]}$$

Finally we observe the easily verified identity

$$x_0 \frac{1}{(1-x_1 E)} y_0 = \frac{1}{(1-x_0 E)} x_0 y_0 . \qquad \text{[Eq. 44]}$$

Care has to be exercised in maintaining the proper order of the operators, proceeding systematically from right to left.

Now from Eq. 18 with $N = 1$ and $k = 1$ we have

$$p_0(t) = (1-e^{-\mu t}) A_c(t) + (1-e^{-\mu t}) \int_0^t a(s) p_0(t-s) ds ,$$

which leads upon Laplace transformation to

$$\pi_0 = \nabla \left[f_0 + \alpha_0 \pi_0 \right] . \qquad \text{[Eq. 45]}$$

We write

$$\left.\begin{aligned} F_k &= \alpha_0\ \pi_k \\ G_0 &= f_0 + F_0 . \end{aligned}\right\} \qquad \text{[Eq. 46]}$$

Then from Eq. 45

$$G_0 = f_0 + \alpha_0\ \nabla G_0$$

or

$$(1 - \alpha_0 \nabla)\ G_0 = f_0 .$$

Replacing ∇ by $1 - E$ gives

$$(\delta_0 + \alpha_0 E)\ G_0 = f_0 ,$$

or

$$\left(1 + \frac{\alpha_0}{\delta_0}\ E\right) G_0 = \frac{f_0}{\delta_0} ,$$

i.e.

$$G_0 = \left(1 + \frac{\alpha_0}{\delta_0}\ E\right)^{-1} \frac{f_0}{\delta_0} . \qquad \text{[Eq. 47]}$$

Next we turn to the formula for $p_1(t)$. This is

$$p_1(t) = e^{-\mu t} A_c(t) + (1 - e^{-\mu t}) \int_0^t a(t-s) p_1(s) + e^{-\mu t} \int_0^t a(t-s) p_0(s)\ ds ,$$

which gives

$$\pi_1 = E\,G_0 + \nabla F_1$$

or

$$F_1 = \alpha_0\, E\, G_0 + \alpha_0\, \nabla F_1 \,, \qquad \text{[Eq. 48]}$$

and hence

$$F_1 = \left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-1} \left(\frac{\alpha_0}{\delta_0} E\, G_0\right) . \qquad \text{[Eq. 49]}$$

From Eq. 47

$$E\,G_0 = \left(1 + \frac{\alpha_1}{\delta_1} E\right)^{-1} \frac{f_1}{\delta_1} \,,$$

so that

$$F_1 = \left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-1} \left[\frac{\alpha_0}{\delta_0} \left(1 + \frac{\alpha_1}{\delta_1} E\right)^{-1} \frac{f_1}{\delta_1}\right]$$

$$= \left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-1} \left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-1} \frac{\alpha_0\, f_1}{\delta_0\, \delta_1}$$

after using an identity of the form of Eq. 44. Thus symbolically

$$F_1 = \left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-2} \left(\frac{\alpha_0\, f_1}{\delta_0\, \delta_1}\right) . \qquad \text{[Eq. 50]}$$

For $k \geq 2$,

$$p_k(t) = e^{-\mu t} \int_0^t a(t-s)\; p_{k-1}(s) ds + (1-e^{-\mu t}) \int_0^t a(t-s)\; p_k(s)\, ds \,, \qquad \text{[Eq. 51]}$$

which leads to

$$F_k = \left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-1} \left[\frac{\alpha_0}{\delta_0} E \left(F_{k-1}\right)\right]. \qquad \text{[Eq. 52]}$$

We are naturally led to conjecture that

$$F_k = \left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-(k+1)} \left[\frac{\alpha_0 \alpha_1 \cdots \alpha_{k-1} \, f_k}{\delta_0 \delta_1 \cdots \delta_k}\right]. \qquad \text{[Eq. 53]}$$

Then

$$EF_{k-1} = \left(1 + \frac{\alpha_1}{\delta_1} E\right)^{-k} \left[\frac{\alpha_1 \alpha_2 \cdots \alpha_{k-1} \, f_k}{\delta_1 \delta_2 \cdots \delta_k}\right].$$

So that the right hand side of Eq. 42 becomes

$$\left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-1} \left[\left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-k} \left(\frac{\alpha_0 \alpha_1 \cdots \alpha_{k-1} \, f_k}{\delta_0 \delta_1 \cdots \delta_k}\right)\right]$$

after using an identity of the type of Eq. 44, which can easily be verified. This confirms Eq. 53 for $k \geq 1$ and leads directly to Eq. 39 for $k \geq 1$ upon expansion and correct interpretation of the binomial operator. From Eq. 47

$$\alpha_0 \pi_0 = F_0 = -f_0 + \left(1 + \frac{\alpha_0}{\delta_0} E\right)^{-1} \frac{f_0}{\delta_0}$$

$$= -f_0 + \frac{f_0}{\delta_0} - \frac{\alpha_0 \, \alpha_1}{\delta_0 \, \delta_1} + \frac{\alpha_0 \alpha_1 \, f_2}{\delta_0 \delta_1 \delta_2} \cdots\cdots,$$

whence

$$\pi_0 = \frac{f_0}{\delta_0} - \frac{f_1}{\delta_0 \delta_1} + \frac{\delta_1 \, f_2}{\delta_0 \delta_1 \delta_2} \cdots\cdots,$$

which confirms the special form of Eq. 39 for $k = 0$. This result can be cross-checked for the system $M/M/\infty$ by putting $b(t) = \mu e^{-\mu t}$ in Eq. 29a, taking the Laplace transform, and then reproducing the above result for the particular case $a(t) = \lambda e^{-\lambda t}$.

3.2 The Output

The output from $X/M/\infty$ can be dealt with in similar fashion. Let $\eta(t)$ be the number of units processed by the system in $(0,t)$ and recall the definition Eq. 31 with $n = 1$, that

$$v_k(t) = \Pr\left[\eta(t) = k \mid \xi(0) = 1\right] .$$

We use $\Phi_k(z)$ for the Laplace transform of $v_k(t)$ and subsequently omit the argument. In addition we write

$$H_k = \alpha_0 \, \Phi_k \, . \qquad \text{[Eq. 54]}$$

We shall show that

$$H_0 = (1 - \alpha_0 E)^{-1} \left[\alpha_0 f_1\right] \qquad \text{[Eq. 55]}$$

and that for $k \geq 1$

$$H_k = \left[\frac{1}{1 - \alpha_0 E} \cdot \alpha_0 \nabla\right]^k \left[\frac{1}{1 - \alpha_0 E} \; f_0\right] . \qquad \text{[Eq. 56]}$$

This can be more explicitly, but heavily, expressed in various ways. It seems best to leave the formula in symbolic form, though

we shall apply the check of showing that it leads to

$$\sum_{k\geq 0} v_k(t) = 1.$$

From Eq. 35a we have, when $n=1$ and $k=0$,

$$v_0(t) = e^{-\mu t}\left[A_c(t) + \int_0^t a(t-s)\; v_0(s)\; ds\right],$$

which gives

$$\Phi_0 = E\left(f_0 + \alpha_0\; \Phi_0\right), \text{ or}$$

$$H_0 = \alpha_0\; E\left(f_0 + H_0\right).$$

Hence

$$(1 - \alpha_0 E)\, H_0 = \alpha_0\; E\, f_0 = \alpha_0\, f_1\,.$$

This is Eq. 55. For $k=1$ we have

$$v_1(t) = (1-e^{-\mu t})A_c(t)+(1-e^{-\mu t})\int_0^t a(t-s)\,v_0(s)ds+e^{-\mu t}\int_0^t a(t-s)\,v_1(s)\,ds$$

or

$$H_1 = \alpha_0\;\nabla\left[f_0 + H_0\right] + \alpha_0\; E(H_1)\,.$$

Now

$$f_0+H_0 = \left[1 + \frac{1}{1-\alpha_0 E}\cdot \alpha_0 E\right]\; f_0 = \frac{1}{1-\alpha_0 E}\cdot f_0\;.$$

Hence

$$H_1 = \frac{1}{1 - \alpha_0 E} \cdot \alpha_0 \nabla \left[\frac{1}{1 - \alpha_0 E} f_0\right],$$

which is Eq. 58 with $k = 1$. In general

$$v_k(t) = (1-e^{-\mu t}) \int_0^t a(t-s) v_{k-1}(s) ds + e^{-\mu t} \int_0^t a(t-s) v_k(s)\, ds,$$

or

$$(1 - \alpha_0 E)\, H_k = \alpha_0 \nabla\, H_{k-1} ,$$

which plainly leads to Eq. 56.

As a check we have:

$$\sum_{k \geq 0} H_k = H_0 + \frac{1}{1 - \alpha_0 E} \cdot \alpha_0 \nabla \cdot \frac{1}{1 - \frac{1}{1 - \alpha_0 E} \cdot \alpha_0 \nabla} \left[\frac{1}{1 - \alpha_0 E} f_0\right]$$

$$= H_0 + \frac{1}{1 - \alpha_0 E} \alpha_0 \nabla \left[\frac{f_0}{\delta_0}\right],$$

since $\nabla + E = 1$.

Hence

$$\sum_{k\geq 0} H_k = \frac{1}{1-\alpha_0 E} \cdot \alpha_0 \left[f_1 + \frac{f_0}{\delta_0} - \frac{f_1}{\delta_1} \right] = \frac{1}{1-\alpha_0 E} \cdot \alpha_0 \left[\frac{f_0}{\delta_0} - \alpha_1 \frac{f_1}{\delta_1} \right]$$

$$= \frac{1}{1-\alpha_0 E} \cdot \alpha_0 \left[1 - \alpha_1 E \right] \frac{f_0}{\delta_0}$$

$$= \frac{1}{1-\alpha_0 E} \cdot \left[1 - \alpha_0 E \right] \cdot \frac{\alpha_0 f_0}{\delta_0}$$

$$= \frac{\alpha_0}{z} \; ,$$

since $f_0 = \delta_0/z$.

Recalling Eq. 54 this means that

$$\sum_{k\geq 0} v_k(t) = 1 \; ,$$

as it should.

The author believes that the results in this chapter have not been published before.

4. THE THEORY OF A CONSTANT ATTRITION ANTISUBMARINE BARRIER

Description in Terms of a Single Server Queueing System in which the Rate of Service depends on the Number in the Queue.

4.1 Introductory

An antisubmarine model to which the theory given here is applicable concerns the situation in which all enemy submarines proceeding from base to operational area are constrained by geographical or other considerations to transit a region in which an antisubmarine barrier is placed. It is supposed that the barrier can exercise constant destructive power per transit. This means that if the enemy force contains n submarines at time t, the probability of a submarine loss in the small time interval $(t, t+dt)$ is $\mu n dt$. This factor coupled with the constant probability differential λ of new construction accounts for the fluctuations in n. The same model can be used to describe air defence situations (raiders attacked by missile defence).

A novel application of this model in the non-military field of congestion theory will be defined in the next section and the analysis will be prosecuted in terms of this model. The terminology is not, in fact, of great importance, for it is the formal similarity of the general model to that of an infinite service system that permits one to borrow from that theory, thereby achieving a simplicity in development that would be difficult to obtain by direct methods. In Sect. 4.3 we will return to a discussion of the antisubmarine model.

Research in the theory of congestion is at present embodied in several books and perhaps about 1500 research papers, the

majority of which have been published since 1950. With relatively few exceptions these studies have been made of systems in which the service and arrival processes are independent. Such systems suffer from the operational defect that when the traffic intensity, measured by the ratio of mean arrival rate to mean service rate, approaches unity the queue tends to elongate and waiting times of the clients become excessive. On the other hand, at low traffic intensities the service facility tends to suffer long periods of idleness. The obvious remedy for this unsatisfactory state of affairs, which seems to be implemented in practice, is to arrange, as far as possible, for service to be matched to demand. (An alternative might be to match demand to service.) Yet, curiously enough, an almost negligible amount of effort has been expended on assessing the effect of this by analyzing appropriate models. This is all the more curious at the present time, in that the effective operation of computer time-sharing systems may be said to depend on the implementation of systems of this nature.

4.2 Fundamental Model

A series of papers by the author and Hadidi [Refs. 7 to 10] have expounded the theoretical and numerical analysis of a system in which the allocation of service time to a given customer is determined by the interval of time between the epochs of arrival of that customer and his predecessor; substantial practical system improvements are demonstrated. The same authors have in addition conducted research on another system in which the instantaneous rate of service depends on the number in the queue [Refs. 11 and 12].

As we shall now explain, the analysis of this latter model is facilitated by analogy with an infinite capacity service facility.

The system under consideration can be explained as follows. A simple $M/M/1$ system has a single server and is characterized by independent negative exponential distributions of arrival and service with mean rates λ and μ respectively. We suppose

now that the service parameter is σ_n when there are n in the queue and in particular that

$$\sigma_n = (n+1)\mu .$$

This model then supposes that the length of time required to serve a given customer depends probabilistically on the number of arrivals during his service time. A large number waiting is characterized by rapid service. In what follows we shall further suppose that arrivals form a Poisson stream with parameter λ and that there is just one server.

Let $\xi(t)$ be the number present in the system (i.e. in the queue and in service) at time t, and write

$$p_n(t) = \Pr\left[\xi(t) = n\right]. \qquad \text{[Eq. 57]}$$

By a familiar argument, which links $\xi(t+dt)$ with $\xi(t)$ via the events that can happen in the small time interval $(t, t+dt)$, we find that $p_n(t)$ is described by the set of equations

$$\left.\begin{array}{l} \left.\begin{array}{l} \dot{p}_0(t) + \lambda p_0(t) = \mu p_1(t) \\ \dot{p}_1(t) + (\lambda+\mu)\, p_1(t) = \lambda p_0(t) + 2\mu p_2(t) \end{array}\right\} \\ \\ \dot{p}_n(t) + (\lambda + n\mu)\, p_n(t) = \lambda p_{n-1}(t) + (n+1)\, \mu p_{n+1}(t) \end{array}\right\} \qquad \text{[Eq. 58]}$$

where dots denote differentiation with respect to time. It will be recognized that these equations are identical with the

set that describes $p_n(t)$ for the infinite service system $M/M/\infty$. Thus $p_n(t)$ has the same value for both systems. There is no question of multiplicative constants, since in both cases $\sum_n p_n(t) = 1$. The results given in Sect. 2.1 for Erlang's problem with $b(t) = \mu e^{-\mu t}$ are therefore applicable.

We shall now show how to deal with the problem of the waiting time of a customer who arrives when the system has been in operation long enough to have "settled down" to a state of statistical equilibrium. This means that the number in the system is described by probabilities $\bar{p}_n$, which are equal to $\lim_{t \to \infty} p_n(t)$ and have the value

$$\bar{p}_n = e^{-\rho} \frac{\rho^n}{n!} . \qquad \text{[Eq. 59]}$$

If we define waiting time to include the customer's own service time and specify a first-come, first-served, queue discipline, then it follows that a customer who arrives to find n already in the system has to wait during the residual service time of the customer being served, plus the service times of the $n-1$ in front of him, plus his own service time. The fact that the service time distribution is composed of negative exponentially distributed components means that residual service time has the same distribution as complete service time. Thus the new customer has to wait for a time that has the distribution of the sum of $n+1$ complete service times.

This leads us to enquire into the form of the joint probability and density $\gamma_{nm}(t)$ that, starting with n, it takes time t to output m successive customers. The waiting time of the customer who arrives to find n in the system is then governed by $\gamma_{n+1,n+1}(t)$, and the steady-state density function $h(w)$

of the waiting time w of a new arrival, however many he finds, is

$$h(w) = e^{-\rho} \sum_{n \geq 0} \frac{\rho^n}{n!} \gamma_{n+1,\ n+1}(w) . \qquad \text{[Eq. 60]}$$

We now proceed to find $\gamma_{nm}(t)$ by invoking the fact that the $\sigma_n = (n+1)\mu$ system is formally identical with $M/M/\infty$. In this infinite system $\gamma_{nm}(t)$ is to be interpreted as referring to the time t required to output m customers, no matter in which order, when there are initially n present. If it is assumed that at $t=0$, service begins on all n customers (the shop opens), the problem can be solved for the system $M/Y/\infty$ in a fashion similar to that employed in Sect. 2.2. This we shall do. We recall that $\eta(t)$ is the number output from $M/Y/\infty$ in $(0, t)$ and that, by Eq. 31,

$$v_k^{(n)}(t) = \Pr \left[\eta(t) = k \mid \xi(0) = n \right] . \qquad \text{[Eq. 61]}$$

Let T_{nm} be the time to the completion of service of the m^{th} customer, given that $\xi(0) = n$. Then evidently,

$$\Pr \left[T_{nm} > t \right] = \Pr \left[\eta(t) < m \mid \xi(0) = n \right]$$

$$= \sum_{k=0}^{m-1} v_k^{(n)}(t) . \qquad \text{[Eq. 62]}$$

Equation 62 determines $\gamma_{nm}(t)$, since

$$\Pr \left[T_{nm} > t \right] = \int_t^{\infty} \gamma_{nm}(s)\, ds.$$

From Eqs. 37 and 4 we have, for $k \le n$,

$$v_k^{(n)}(t) = e^{-\lambda f(t)} \sum_{i=0}^{k} \binom{n}{i} B^i(t)\, B_c^{n-i}(t)\, \frac{\left[\lambda f(t)\right]^{k-i}}{(k-i)!}$$

$$= e^{-\lambda f(t)}\, C_k^{(x)}\left\langle e^{\lambda x f(t)}\left[B_c(t) + x\, B(t)\right]^n\right\rangle , \qquad \text{[Eq. 63]}$$

where $C_k^{(x)}\left\langle \quad \right\rangle$ means the coefficient of x^k in $\left\langle \quad \right\rangle$.

Thus

$$v_k^{(n)}(t) = \frac{e^{-\lambda f(t)}}{2\pi i} \int_\Gamma e^{\lambda z f(t)} \left[B_c(t) + z\, B(t)\right]^n \frac{dz}{z^{k+1}} ,$$

where Γ is a contour in the z plane enclosing $z = 0$. Then

$$\Pr\left[T_{nm} > t\right] = \frac{e^{-\lambda f(t)}}{2\pi i} \int_\Gamma e^{\lambda z f(t)} \frac{\left[B_c(t) + z\, B(t)\right]^n}{(1 - z)} \frac{dz}{z^m} , \qquad \text{[Eq. 64]}$$

provided that Γ excludes $z = 1$. Hence

$$H_c(t) = \int_t^\infty h(s)ds = e^{-\rho} \sum_{n \ge 1} \frac{\rho^{n-1}}{(n-1)!} \Pr\left[T_{nn} > t\right]$$

$$= \frac{e^{-\rho-\lambda f(t)}}{2\pi i} \int_\Gamma \exp\left\langle \lambda z f(t) + \rho\left[B_c(t) + zB(t)\right] / z \right\rangle \frac{\left[B_c(t) + z\, B(t)\right] dz}{z(1 - z)} . \qquad \text{[Eq. 65]}$$

Equation 65 gives the tail of the waiting time distribution for the single server "$\sigma_n = (n+1)\mu$" queueing system when $b(t)$ is replaced by $\mu e^{-\mu t}$. The result is delivered much more expeditiously by utilizing the infinite service system analogy than by any direct approach that we have seen. [See Hadidi (Ref. 12) where Eq. 67 is obtained by complicated analytical methods.]

$H_c(t)$ can be expressed explicitly in terms of modified Bessel functions of the first kind by expressing Eq. 65 in terms of the well-known contour integral representations of the latter. Thus one obtains

$$H_c(t) = e^{-\xi-\eta}\left[B_c(t)\ I_0(2\sqrt{\xi\eta}) + \sum_{n\geq 0} \left(\frac{\eta}{\xi}\right)^{\frac{1}{2}(n+1)} I_{n+1}(2\sqrt{\xi\eta})\right], \qquad \text{[Eq. 66]}$$

where

$$\xi = \lambda f(t)$$

$$\eta = \rho B_c(t).$$

Similarly, the density function is given by

$$h(t) = e^{-\xi-\eta}\left\{\left[\lambda B(t)B_c(t) + b(t)(1+\rho B(t))\right] I_0(2\sqrt{\xi\eta}) + \right.$$

$$\left. + \left[\rho b(t)B_c(t)\left(\frac{\xi}{\eta}\right)^{\frac{1}{2}} + \lambda B^2(t)\left(\frac{\eta}{\xi}\right)^{\frac{1}{2}}\right] I_1(2\sqrt{\xi\eta})\right\}, \qquad \text{[Eq. 67]}$$

which may be evaluated either directly from the contour integral obtained from Eq. 65 by differentiation, or by differentiating Eq. 66 and utilizing known Bessel function relations.

Equation 65 can be checked by setting $t = 0$. $H_c(0)$ should have the value unity. Now $f(0) = 0$, $B_c(0) = 1$, $B(0) = 0$. Hence

$$H_c(0) = \frac{e^{-\rho}}{2\pi i} \int_{\Gamma} e^{\rho/z} \frac{dz}{z(1-z)}$$

$$= \frac{1}{2\pi i} \int_{\Gamma'} e^{\rho\xi} \frac{d\xi}{\xi} ,$$

by an obvious transformation.

This is clearly unity. Equation 66 may also be checked directly in the same way.

An explicit treatment of some analytical and practical aspects of the $\sigma_n = (n+1)\mu$ single-server system can be found in Hadidi and Conolly [Ref. 11].

4.3 ASW Application

The interpretation of this result will now be made in the context of the antisubmarine model referred to in Sect. 4.1, in which all enemy submarines proceeding to and from their operational area are obliged to transit an antisubmarine barrier that has constant destructive power per transit.

The commander who asks how long it will take to destroy as many submarines as the enemy now possesses can be answered by using the fact that the distribution of the time is $1 - H_c(t)$. This is the interpretation of the single-server waiting time result. Note that unless λ is reduced to zero the enemy is still likely to possess submarines even when a number equal to his present strength has been destroyed.

A cognate question is: "How long before I first reduce the enemy's present strength to zero?". This is of interest to both offence and defence. More generally it can be framed in terms of a reduction from the present level to some other prescribed non-zero level. Such questions are answered in the terminology of stochastic processes by consideration of the distributions of "first passage times". In congestion theory the interval of time from the epoch at which one client first occupies the system until for the first time thereafter it becomes empty is described as a "busy period", and one says that in this time the system has made a first passage from unity to zero. Let $k(t)$ be the density function of a busy period for a system in which arrivals and services have negative exponential characteristics. Then the time required for a first passage from N to n $(N > n)$ has density function $k^{(N-n)}(t)$, the bracket superscript denoting convolution. A congestion model with negative exponential arrivals corresponds to an antisubmarine barrier model with time-independent building rate.

The writer has analyzed the problem of the busy period for the $M/Y/\infty$ congestion system, and by analogy this is applicable with negative exponential service time to the barrier model in the sense mentioned above. The result may be expressed as follows.

Let

$$b(t) = \mu e^{-\mu t},$$

$$B(t) = \int_0^t b(s)\, ds,$$

$$f(t) = \int_0^t B(s)\, ds.$$

Let

$$h(x, t) = e^{xf(t)} - 1,$$

where x is some real number, and let $h^{(n)}(x, t)$ be the n-fold convolution of $h(x, t)$ with itself (with respect to t).

Moreover, let

$$h^{(n)}(x, t) = \sum_{m \geq n} x^m h_{nm}(t).$$

Then if $D^n \equiv \frac{d^n}{dt^n}$

$$k(t) = \frac{e^{-\lambda t}}{\lambda} \sum_{n \geq 0} (-1)^n D^{n+2} \left[h^{(n+1)} (1, t) \right]$$

This result appears to be new. Its proof is given in Appendix A.

In terms of the model of units in a geographical area the functions $\gamma_{nm}(t)$ give information on the time required to output m units, given that n are present at a certain moment. If the units are merchant ships carrying strategic material such information seems logistically pertinent.

REFERENCES

1. Brockmeyer, Halstrøm and Jensen, "The Life and Works of A.K. Erlang", The Copenhagen Telephone Company, 1948.

2. A.Y. Khintchine, "Mathematical Methods in the Theory of Queueing", Charles Griffin, London. (Translation of earlier work, including "Mathematical Theory of a Stationary Queue". Mat. Sbornik, Vol. 39, No. 4, pp.103-144, 1932). (Translation 1960).

3. L. Takács, "Introduction to the Theory of Queues", Oxford University Press, 1932.

4. N.M. Mirasol, "The Output of an $M/G/\infty$ Queueing System is Poisson", Operations Research, Vol. 11, pp.282-284, 1963.

5. D.G. Kendall, "Some Recent Work and further Problems in the Theory of Queues", Theory of Probability and its Applications", Vol. IX, No. 1, pp.1-13, 1964.

6. D. Vere-Jones, "Some Applications of Probability Generating Functionals to the Study of Input-output Streams", Journal of Royal Statistical Society, Vol. 30, No. 2, pp.321-333, 1968.

7. B.W. Conolly, "The Waiting Time Process for a certain Correlated Queue", Operations Research: Journal Operations Research Society of America, Vol. 16, No. 5, pp. 1006-1015, 1968.

8. B.W. Conolly and N. Hadidi, "A Correlated Queue", to be published in Journal Applied Probability, 1969.

9. B.W. Conolly and N. Hadidi, "A Comparison of the Operational Features of the Conventional M/M/1, D/M/1 and M/D/1 Queues with those of a Self-regulating System", to be published in Applied Statistics, 1969.

10. N. Hadidi, "Correlation between Arrival and Service Patterns as a Means of Queue Regulation", Doctoral dissertation. Department of Statistics, Virginia Polytechnic Institute, 1968.

11. N. Hadidi and B.W. Conolly, "On the Improvement of the Operational Characteristics of Single Server Queues by the Use of a Queue Length Dependent Mechanism", to be published in Applied Statistics, 1969.

12. N. Hadidi, "On the Service Time Distribution and the Waiting Time Process of a Potentially Infinite Capacity Queueing System", to be published in the Journal of Applied Probability, 1969.

APPENDIX A

THE BUSY PERIOD FOR THE INFINITE CAPACITY SERVICE SYSTEM $M/G/\infty$

In the terminology of Queuing Theory a busy period (B.P.) begins when the service facility ceases to be idle, and it continues until, for the first time thereafter, it becomes idle again. An infinite capacity system is supposed to be supplied with an infinite number of servers. Immediately before a B.P. they are all idle. At the beginning of the B.P. an arrival occurs and immediately begins to be served. The B.P. continues until once again all servers are idle. It is hardly necessary to add that since the number of servers is supposed infinite not all servers will be continuously occupied during a B.P. as here defined. Let $\eta(T)$ be the number of customers served during a B.P. that begins at time $t = 0$ and terminates at $t = T$. We are primarily concerned in this paper with the joint probability

$$k_n(t)dt = \Pr\Big[t < T < t+dt,\ \eta(T) = n\Big] . \qquad \text{[Eq. A.1]}$$

The $M/G/\infty$ system is characterized by an arrival pattern that forms a Poisson stream with time-independent parameter λ. Thus interarrival intervals have p.d.f. $\lambda e^{-\lambda t}$. A service period initiated by any one of the infinite number of servers is supposed to have a proper p.d.f. $b(t)$ and a distribution function denoted by $B(t)$, such that $B(\infty) = 1$. The duration of a service is independent of the arrival stream.

We also introduce the function

$$f(t) = \int_0^t B(s)\, ds, \qquad \text{[Eq. A.2]}$$

and a generating function $h(x, t)$ defined by

$$h(x, t) = \exp\left[\lambda x f(t)\right] - 1 , \qquad \text{[Eq. A.3]}$$

where x can be taken to be a real variable restricted as necessary to ensure convergence. The commonly accepted notation $p(t)*q(t)$ will be used for the convolution of two functions, viz.

$$p(t)*q(t) = \int_0^t p(s)\, q(t-s)\, ds,$$

and $p^{(n)}(t)$ will denote the n-fold convolution of $p(t)$ with itself, viz.

$$p^{(n)}(t) = p^{(n-1)}(t) * p(t),$$

with the convention that $p^{(1)}(t) \equiv p(t)$. Then convolution of $h(x, t)$ with itself and with respect to t produces a higher order generating function $h^{(n)}(x, t)$, which, when expanded in ascending powers of x has coefficients that will be denoted by $h_{nm}(t)$. Thus

$$h^{(n)}(x, t) = \sum_{m \geq n} h_{nm}(t)\, x^m, \qquad \text{[Eq. A.4a]}$$

and we observe then from Eq. A.3 that

$$h_{1m}(t) = \left[\lambda f(t)\right]^m / m! \; . \qquad \text{[Eq. A.4b]}$$

In the following development it will simplify writing to replace expressions like $d^n p(t)/dt^n$ by $D^n p(t)$. In general when there is no ambiguity we shall omit arguments.

The objective of the paper is to prove the result

$$k_n(t) = \frac{e^{-\lambda t}}{\lambda} \left[D^2(h_{1n}) - D^3(h_{2n}) + \ldots\ldots + (-)^{n+1} D^{n+1}(h_{nn}) \right] \quad \text{[Eq. A.5]}$$

The proof will proceed on the following lines. Suppose that customers leaving the infinite service system are labelled in the order of their departure (1, 2, 3, ... n, ...). This sequence is a permutation of their order of arrival. We suppose that at $t = 0$ a B.P. begins and consider the event E_n that the customer labelled n has been served in $(t, t + dt)$ leaving the system empty. There are two mutually exclusive ways in which E_n can occur: either it is the termination of the busy period beginning at $t = 0$; or there may intervene one or more idle periods before t. We evaluate the unconditional probability of occurrence of E_n, and the probability of E_n given the intervention of one or more idle periods; $k_n(t)dt$ is then the difference between these quantities.

We first require two lemmata.

Lemma 1. It is supposed that the system is empty at $t = 0$. Let $r_n(t)$ be the probability that exactly n arrivals take place during $(0, t)$ and that all of them, and no more, have been served before t. Then

$$r_n(t) = e^{-\lambda t} h_{1n}. \quad \text{[Eq. A.6]}$$

Proof. Consider first $r_1(t)$. Let the first and only arrival occur at epoch u. He has to have been served before the

termination of the remainder of the interval, which has length $t - u$. Then

$$r_1(t) = \lambda \int_0^t e^{-\lambda u} e^{-\lambda(t-u)} B(t-u)\, du$$

$$= e^{-\lambda t} f(t) = e^{-\lambda t} h_{1,1}(t) . \qquad \text{[Eq. A.7]}$$

Now suppose that Eq. A.6 holds for all n, and consider $r_{n+1}(t)$. If again the first arrival occurs at epoch u, then during $t - u$

(a) this customer must have been served,

(b) independently, n further customers have arrived and been served.

The fact that we have an infinite system enables us to make independent arrangements for the first arrival and to separate his channel from the system as long as it is occupied. Thus at epoch u the initial conditions of the lemma are reproduced and we have

$$r_{n+1}(t) = \lambda \int_0^t e^{-\lambda u} B(t-u)\, r_n(t-u)\, du$$

$$= \frac{\lambda^{n+1}}{n!} \int_0^t e^{-\lambda u} B(t-u)\, e^{-\lambda(t-u)} f^n(t-u)\, du$$

by Eqs. A.6 and A.4b

$$= \frac{\lambda^{n+1}}{n!} \quad e^{-\lambda t} \int_0^t f^n \; D(f) \; du$$

$$= e^{-\lambda t} \; \frac{(\lambda f)^{n+1}}{(n+1)!}$$

$$= e^{-\lambda t} \; h_{1,n+1} \; .$$

This completes the proof of Eq. A.6.

Lemma 2. Let $t=0$ be the beginning of a B.P. Let T_n be the epoch at which the n^{th} service was completed, not necessarily in order of arrival and not necessarily without the intervention of one or more idle periods in the service facility, and with an empty system immediately after T_n. Let

$$v_n(t) \; dt \; = \; Pr(t < T_n < t + dt).$$

Then

$$v_n(t) = \frac{e^{-\lambda t}}{\lambda} \; D^2 \; (h_{1,n}) \; . \qquad \text{[Eq. A.8]}$$

(Note that, $v_n(t)$ is the p.d.f. of the event E_n referred to earlier.

Proof. This is again inductive. To deal with $v_1(t)$ we require no arrivals in $(0, t)$ and the completion of service

of the original customer in $(t, t+dt)$. Thus

$$v_1(t) = e^{-\lambda t} b(t)$$

$$= e^{-\lambda t} D^2(f)$$

$$= \frac{e^{-\lambda t}}{\lambda} D^2 (h_{1;1}) . \qquad \text{[Eq. A.9]}$$

Now consider v_{n+1}. If it is the original customer who has been served in $(t, t+dt)$ we have a contribution of amount

$$b(t)\, r_n(t) = D(B)\, e^{-\lambda t}\, h_{1n}$$

to v_{n+1}, by Eq. A.6 and the separation of channels concept used in the proof of Lemma 1. If the original customer has been served before t [probability $B(t)$] we bring in a first arrival at epoch $t = u$, which, since the original arrival passes independently through the system, reproduces the condition at $t = 0$. Thus

$$v_{n+1}(t) = e^{-\lambda t} D(B)\, h_{1n} + \lambda B \int_0^t e^{-\lambda(t-u)}\, v_n(u)\, du. \qquad \text{[Eq. A.10]}$$

By Eq. A.8 this gives

$$v_{n+1}(t) = e^{-\lambda t} \left[h_{1n}\, D(B) + B\, D(h_{1n}) \right]$$

$$= e^{-\lambda t} D(h_{1n}\, B)$$

$$= \frac{\lambda^n}{n!} e^{-\lambda t} D(f^n B)$$

$$= \frac{\lambda^n}{(n+1)!} e^{-\lambda t} D(f^n B)$$

$$= \frac{e^{-\lambda t}}{\lambda} D^2 (h_{1,n+1}) .$$

This proves Eq. A.8.

To prove the main result [Eq. A.5] we again suppose the conditions of Lemma 2 and, recalling earlier statements, observe that

$$\Pr\left[t < T_n < t+dt\right] = \Pr\left[t < T_n < t+dt, \begin{array}{l}\text{no idle periods}\\ \text{before } t\end{array}\right]$$

$$+ \Pr\left[t < T_n < t+dt, \begin{array}{l}\text{at least one idle}\\ \text{period before } t\end{array}\right];$$

[Eq. A.11]

i.e. that

$$v_n(t) = k_n(t) + \ell_n(t), \qquad \text{[Eq. A.12]}$$

where $\ell_n(t)$ refers to the second term on the right-hand side of Eq. A.11. Knowing v_n, if we can determine ℓ_n, then we shall have obtained k_n. Suppose the first idle period occurs after the termination of the m^{th} service $(1 \le m \le n-1)$ at epoch t-u. Then (0, t-u) is a B.P. with m services. We then have to arrange for a further arrival at epoch $t-u+\omega(0<\omega<u)$, which reproduces the initial conditions.

During $u-\omega$, $n-m$ services are required, with or without idle periods, the last terminating at epoch t. Hence

$$\ell_n(t) = \lambda \sum_{m=1}^{n-1} \int_0^t k_m(t-u)du \int_0^u e^{-\lambda\omega} v_{n-m}(u-\omega)\, d\omega$$

$$= \sum_{m=1}^{n-1} k_m(t) * \quad e^{-\lambda t} D\left[h_{1n-m}(t)\right] \quad ,$$

using Eq. A.8. Now using Eq. A.5 for $1 \le m \le n-1$, we have

$$\ell_n(t) = \frac{e^{-\lambda t}}{\lambda} \left[D^3\left(h_{1,1} * h_{1n-1}\right) + \right.$$

$$+ D^3\left(h_{1,2} * h_{1n-2}\right) - D^4\left(h_{2,2} * h_{1n-2}\right) +$$

$$+ D^3\left(h_{1,3} * h_{1n-3}\right) - D^4\left(h_{2,3} * h_{1n-3}\right) + D^5\left(h_{3,3} * h_{1n-3}\right)$$

$$+ \ldots\ldots\ldots\ldots$$

$$+ D^3\left(h_{1n-1} * h_{1,1}\right) - D^4\left(h_{2n-1} * h_{1,1}\right) + D^5\left(h_{3n-1} * h_{1,1}\right) -$$

$$\ldots\ldots + (-)^n D^{n+1}(h_{n-1,n-1} * h_{1,1})$$

$$= \frac{e^{-\lambda t}}{\lambda} \left[D^3(h_{2n}) - D^4(h_{3n}) + D^5(h_{4n}) - \ldots + (-)^n D^{n+1}(h_{nn})\right.$$

[Eq. A.13]

This, with Eq. A.12 establishes Eq. A.5, it remaining only to show from first principles that

$$k_1(t) = \frac{e^{-\lambda t}}{\lambda} D^2(h_{1,1}), \qquad \text{[Eq. A.14]}$$

which is obvious since $D^2(h_{1,1}) = \lambda b$.

The p.d.f. $k(t)$ of a busy period of length t is given by

$$k(t) = \sum_{n\geq 1} k_n(t). \qquad \text{[Eq. A.15]}$$

Using Eq. A.5 we obtain

$$k(t) = \frac{e^{-\lambda t}}{\lambda} \sum_{n\geq 0} (-)^n D^{n+2} \left[h^{(n+1)}(t)\right], \qquad \text{[Eq. A.16]}$$

where $h^{(n)}(t) = h^{(n)}(1, t)$.

As a check we now propose to obtain Eq. A.16 by alternative means. The event F consisting of the beginning of a B.P. is recurrent for $M/G/\infty$ and the probability of its occurrence in $(t, t+dt)$ is $\lambda p_0(t)dt$, where $p_0(t)$ is the probability that a system intiated at $t=0$ with one member will be empty at time t. The interval between two successive occurrences has density function $\lambda k(t) * e^{-\lambda t}$, and so

$$\lambda p_0(t) = \lambda k * e^{-\lambda t} + (\lambda k * e^{-\lambda t}) * \left[\lambda p_0(t)\right], \qquad \text{[Eq. A.17)}$$

since F may occur for the first time in $(t, t+dt)$, or may occur at a previous epoch.

If $\pi_0(z)$ and $K(z)$ are the Laplace transforms of $p_0(t)$ and $k(t)$, then Eq. A.17 is equivalent to

$$K(z) = \frac{(\lambda+z)\pi_0(z)}{1+\lambda\pi_0(z)} . \qquad \text{[Eq. A.18]}$$

This gives formally

$$K(z) = \sum_{n\geq 0} (-\lambda)^n (\lambda+z)\, \pi_0{}^{n+1}(z). \qquad \text{[Eq. A.19]}$$

Now it is well known [e.g. Khintchine (Ref. 2)] that

$$p_0(t) = e^{-\lambda t+\lambda f} \quad B = \frac{e^{-\lambda t}}{\lambda} D\left[h(t)\right], \qquad \text{[Eq. A.20]}$$

where $h(t) = h(1, t)$.

If we write $H(z)$ for the Laplace transform of $h(t)$ it follows immediately from Eq. A.20 that

$$\pi_0(z) = (\lambda+z)\, H(z+\lambda)/\lambda$$

and hence that Eq. A.19 gives

$$K(z) = \lambda^{-1} \sum_{n\geq 0} (-)^n (\lambda+z)^{n+2} H^{n+1}(z+\lambda).$$

This transforms back into

$$k(t) = e^{-\lambda t} \lambda^{-1} \sum_{n \geq 0} (-)^n D^{n+2} \left[h^{(n+1)} \right],$$

which is Eq. A.16.

Thus we have a check on the purely probabilistic arguments of the earlier paragraphs. Because of Eq. A.18 and the known properties of $p_0(t)$, we also observe from Eq. A.18 that

$$\int_0^\infty k(t)dt = 1.$$

The mean B.P. length is perhaps most expeditiously obtained from the relation, applicable to $M/G/\infty$, that

$$\frac{E(I)}{E(BP) + E(I)} = \lim_{t \to \infty} p_0(t) = e^{-\rho}, \qquad \text{[Eq. A.21]}$$

where $E(\ )$ means "expected value of", and I denotes an idle period. This gives

$$E(BP) = \frac{1}{\lambda}(e^{\rho} - 1), \qquad \text{[Eq. A.22]}$$

where $\rho = \lambda/E$ (service interval).